



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/593,275	06/13/2000	Upendra V. Chaudhari	YOR-2000-0168US1	7772

7590 02/03/2003

Ference & Associates
129 Oakhurst Road
Pittsburgh, PA 15215

EXAMINER

HAN, QI

ART UNIT

PAPER NUMBER

2654

DATE MAILED: 02/03/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/593,275

Applicant(s)

CHAUDHARI ET AL.

Examiner

Qi Han

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-27 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1-3, 6-12, 14-16, 19-25 and 27 is/are rejected.
- 7) ☒ Claim(s) 4, 5, 13, 17, 18 and 26 is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on ____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) ☐ The proposed drawing correction filed on ____ is: a) ☐ approved b) ☐ disapproved by the Examiner.
If approved, corrected drawings are required in reply to this Office action.
- 12) ☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

- 13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. ____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
* See the attached detailed Office action for a list of the certified copies not received.
- 14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
a) ☐ The translation of the foreign language provisional application has been received.
- 15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892) 4) ☐ Interview Summary (PTO-413) Paper No(s). ____
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948) 5) ☐ Notice of Informal Patent Application (PTO-152)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 5. 6) ☐ Other:

DETAILED ACTION

Information Disclosure Statement

1. The references listed in the Information Disclosure Statement submitted on 11/27/2001 have been considered by the examiner (see attached PTO-1449).

Specification

2. The disclosure is objected to because of the following informalities:
 - a. On page 8, equation (1), the variable " $P(u_t | M(i, j, (i, t)))$ " appears to be " $P(u_t | M(i, j(i, t)))$ ". Appropriate correction is required.
 - b. The word "pickmax" is not a commonly accepted technical term in the art, and is not descriptive as part of the title. Appropriate correction is required. The examiner suggests deleting the word "pickmax" from the title.
3. The specification is objected to as failing to provide proper antecedent basis for the claimed subject matter. See 37 CFR 1.75(d)(1) and MPEP § 608.01(o). Correction of the following is required:

Regarding **claims 1, 14 and 27**, the term of "non-interpolated likelihood value" is unclear, since applicant has failed to provide a particular explanation or definition for the term in the specification and the claims. "non-interpolated likelihood value" is not a commonly accepted technical term in the art. Further clarification or definition of the term is required.

Art Unit: 2654

Regarding **claims 13 and 26**, the expression of "S (U|M)" lacks antecedent basis in the claim; it is unclear that which equation will be applied for this expression in the claim. The examiner interprets that the expression "S (U|M)" corresponds to the equation in claim 5 hereinafter, for both claims 13 and 26.

Drawings

4. The drawings are objected to because of the following informalities:

It appears to miss text label for each associated reference number in Fig. 1.

A proposed drawing correction or corrected drawings are required in reply to the Office action to avoid abandonment of the application. The objection to the drawings will not be held in abeyance.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. Claims 1-3, 6-12, 14-16, 19-25 and 27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Goldenthal et al. (USPN 6,205,424) hereinafter referenced as Goldenthal, in view of Newman et al. (USPN 5,946,654) hereinafter referenced as Newman.

Art Unit: 2654

Regarding **claim 1**, as best understood in view of the objection (see above),

Goldenthal discloses two-staged cohort selection for speaker verification system.

Goldenthal further discloses that:

- a. in a speaker verification system, individuals having known identities supply utterances or speech samples during training sessions, and a temporal sequence of observation vectors (frames) from a sample of speech signals contains a set of acoustic features, the frames of the various individuals can be further processed to create models representing the speech (column 1, lines 35-45), and the model generator can use pattern classification and recognition methodology that is a segment based speech processing in that designated segment can be units of speech, such as phones, or transition from one phone to another (column 4, lines 8-29), which corresponds to the claimed "providing a model corresponding to a target speaker, the model being resolved into at least one frame and at least one level of phonetic detail;"
- b. the claimed identity of an individual can be verified by having the individual utter a prompted sequence of words or spontaneous speech during a testing session (column 1, lines 47-49), which corresponds to the claimed "receiving an identity claim;"
- c. the validation or testing speech signals are analyzed and compared with the prestored observation models corresponding to the "claimed" identity to determine scores, the scores can be expressed as log likelihood scores: $\text{score} = \log p(O/I)$, where p represents the likelihood that the observed frames O were produced by the individual I , and if the scores exceed a predetermined threshold, it is presumed that the individual is who he or she claims to be (column 1, lines 50-57), which corresponds to the claimed

“ascertaining whether the identity claim corresponds to the target speaker model; said ascertaining step comprising the steps of: determining, for each frame and each level of phonetic detail of the target speaker model, a non-interpolated likelihood value; and resolving the at least one likelihood value to obtain a likelihood score.”

But, Goldenthal fails to expressly disclose “each level of phonetic detail of target speaker model.” However, the examiner contends that the concept of providing phonetic detail levels was well known, as taught by Newman.

In the same field of endeavor, Newman discloses speaker identification using unsupervised speech models. Newman further discloses that each word 700 (Fig. 7) is represented by a set of phonemes 705 that represent the phonetic spelling of the word, and each phoneme is represented by three sets of model parameters 710 that correspond to the three nodes of the phoneme (column 6, lines 29-34).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Goldenthal by specifically providing phonetic detail levels, as taught by Newman, for the purpose of increasing efficiency and quality of a recognition system.

Regarding **claim 2**, as best understood in view of the objection (see above), Goldenthal and Newman disclose everything claimed, as applied above (see claim 1). Goldenthal further discloses a log likelihood score (column 1, lines 53-54). Moreover, Goldenthal teaches that the normalized score comprises a log likelihood function f that can be statistical in nature, as maximum (column 2, lines 21-31), which suggests that the log likelihood score may be used for

a maximum likelihood calculation, which corresponds to the claimed “for each frame and each level of phonetic detail, the non-interpolated likelihood value is a maximum likelihood value.”

Regarding **claim 3**, Goldenthal and Newman disclose everything claimed, as applied above (see claim 2). Goldenthal further discloses a log likelihood score (column 1, lines 53-54). Moreover, Goldenthal teaches that the normalized score comprises a log likelihood function f that can be statistical in nature, as average (column 2, lines 21-31), which corresponds to the claimed “said step of resolving the at least one likelihood value comprises averaging the at least one likelihood value.”

Regarding **claim 6**, Goldenthal and Newman disclose everything claimed, as applied above (see claim 2). Goldenthal discloses that the model generator can use pattern classification and recognition methodology that is a segment based speech processing in that designated segment can be units of speech, such as phones, or transition from one phone to another (column 4, lines 8-29), which corresponds to the claimed “the at least one level of phonetic detail comprises at least one of the following: a global level; a phonemic level and a subphonemic level.”

Regarding **claim 7**, Goldenthal and Newman disclose everything claimed, as applied above (see claim 6). But, Goldenthal fails to expressly disclose “the at least one level of phonetic detail comprises all of the following three levels: a global level; a phonemic level and a sub-phonemic level.” However, the examiner contends that the concept of providing all phonetic detail levels was well known, as taught by Newman.

Newman further discloses that each word 700 (Fig. 7) is represented by a set of phonemes 705 that represent the phonetic spelling of the word, and each phoneme is

Art Unit: 2654

represented by three sets of model parameters 710 that correspond to the three nodes of the phoneme (column 6, lines 29-34), which may correspond to the claimed “three levels: a global level; a phonemic level and a sub-phonemic level.”

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Goldenthal by specifically providing all phonetic detail levels, as taught by Newman, for the purpose of increasing efficiency and quality of a recognition system.

Regarding **claim 8**, Goldenthal and Newman disclose everything claimed, as applied above (see claim 7). But, Goldenthal fails to expressly disclose “providing labeling information for each frame.” However, the examiner contends that the concept of providing labeling information for each frame was well known, as taught by Newman.

Newman further discloses that in performing speech recognition, the processor processes the sample 110 (Fig 1) to produce a sequence of frames 115 (step 415) and shows that each frame 500 (Fig. 6) of the sequence of frames includes a set of parameters 600 that represent the frequency content of the frame (column 6, lines 11-25). Furthermore, Newman discloses that the relationship or index between a set of digital frames 500 (Figs. 5 and 6) received from speech sample 110 (Fig. 1) and a set of parameters 600 (Fig. 6) representing the content of the frame and a set of words (text) 125 (Figs. 1 and 8) referenced by starting frame number (column 5, line 54, column 6, line 14 and column 8, line 32).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Goldenthal by specifically providing labeling

information for each frame, as taught by Newman, for the purpose of increasing efficiency of a recognition system.

Regarding **claim 9**, Goldenthal and Newman disclose everything claimed, as applied above (see claim 1). Goldenthal further discloses the validation or testing speech signals are analyzed and compared with the prestored observation models corresponding to the "claimed" identity to determine scores and if the scores exceed a predetermined threshold, it is presumed that the individual is who he or she claims to be (column 1, lines 50-57), which corresponds to the claimed "said ascertaining step further comprises accepting or rejecting the identity claim."

Regarding **claim 10**, Goldenthal and Newman disclose everything claimed, as applied above (see claim 9). Goldenthal further discloses that the validation or testing speech signals are analyzed and compared with the prestored observation models corresponding to the "claimed" identity to determine scores and if the scores exceed a predetermined threshold, it is presumed that the individual is who he or she claims to be (column 1, lines 50-57), which corresponds to the claimed "said step of accepting or rejecting comprises comparing a quantity based on the likelihood score to a predetermined threshold value."

Regarding **claim 11**, Goldenthal and Newman disclose everything claimed, as applied above (see claim 10). Goldenthal further discloses that a plurality of sets of "cohort" models (CM) 170 (Fig. 1) that characterize the speech signals of each identified speaker, are selected from the available sets of acoustic models of the other speakers, and the selection can be made according to predetermined selection criteria, for example, the models which best characterize the speech of the identified speaker, or the models whose characterization fits some predetermined probability density function (column 4, lines 49-58), which corresponds to the

claimed “the steps of providing at least one model corresponding to at least one background speaker; and determining the quantity based on the likelihood score via employing the at least one background speaker model.”

Regarding **claim 12**, Goldenthal and Newman disclose everything claimed, as applied above (see claim 11). Goldenthal further discloses that during testing, the score obtained from the models of the speaker whose identity is claimed is compared with all of the scores derived from the small set of cohort models to produce a set of score differences, and the differences are then used as a normalized score $= \log p(O/I) - f[\log p(O/(C_k(I))]$, where $\log p(O/(C_k(I))$ are the scores for the k cohorts linked to the claimed individual (column 2, lines 21-28), which corresponds to the claimed “said step of determining the quantity based on the likelihood comprises determining a log-likelihood ratio based on the likelihood score.”

Regarding **claims 14-16 and 19-25**, they disclose an apparatus, which corresponds to the method of claims 1-3 and 6-12, respectively. The apparatus is obvious in that it simply provides structure for the functionality found in claims 1-3 and 6-12, respectively.

Regarding **claim 27**, it discloses a program storage device readable by machine, which corresponds to the method of claim 1. The program storage device readable by machine is obvious in that it simply provides implementation and structure for the functionality found in claim 1.

Allowable Subject Matter

6. Claims 4-5, 13, 17-18 and 16 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The following is an examiner's statement of reasons for allowance:

Regarding **claim 4**, the prior art of record fail to specifically disclose or fairly suggest a way to determine the likelihood value through a particular equation, as described in the claim, which calculates the likelihood score by using multiple levels of phonetic detail of the speaker model, each level may have multiple processing units, wherein the multiple levels (L) is interpreted as more than one level in most of processing situation.

Regarding **claim 5**, it is dependent claim of the claim 4 and includes all features of its parent claim(s).

Regarding **claim 13**, as best understood in view of the objection (see above), the prior art of record fail to specifically disclose or fairly suggest a way to provides the log-likelihood ration calculation, as described in the claim, employing multiple levels of phonetic detail of the speaker model and background speaker models, wherein each level may have multiple processing units, wherein the multiple levels (L) is interpreted as more than one level in most of processing situation.

Regarding **claim 17,18 and 26**, they are they disclose an apparatus, which corresponds to the method of claims 4, 5 and ¹³26, respectively. The apparatus is inherent in that it simply provides structure for the functionality found in claims 4, 5, and ¹³26, respectively.

Art Unit: 2654

The prior art of record provided numerous teachings of alternating types of speaker recognition, identification and verification. However, the features as presented above are not anticipated by, nor made obvious over the prior art of the record.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Conclusion

7. Any response to this office action should be mailed to:
Commissioner of Patents and Trademarks, Washington D.C. 20231
or faxed to:
(703)-872-9314
Hand-delivered responses should be brought to:
Crystal Park II, 2121 Crystal Drive, Arlington, VA. Sixth Floor
(Receptionist).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Qi Han whose telephone numbers is (703) 305-5631. The examiner can normally be reached on Monday through Thursday from 8:00 a.m. to 5:30 p.m. and Friday from 8:00 a.m. to 12:00 a.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Marsha Banks-Harold, can be reached on (703) 305-4379.

Any inquiry of a general nature of relating to the status of this application or proceeding should be directed to the Technology Center 2600 Customer Service Office whose telephone number is (703) 306-0377.

QH/qh
January 21, 2003

Marsha D Banks-Harold
MARSHA D. BANKS-HAROLD
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600